

Revisiting Complex Moments For 2D Shape Representation and Image Normalization

João B. F. P. Crespo Pedro M. Q. Aguiar

May, 2010

Abstract

When comparing 2D shapes, a key issue is their normalization. Translation and scale are easily taken care of by removing the mean and normalizing the energy. However, defining and computing the orientation of a 2D shape is not so simple. In fact, although for elongated shapes the principal axis can be used to define one of two possible orientations, there is no such tool for general shapes. As we show in the paper, previous approaches fail to compute the orientation of even noiseless observations of simple shapes. We address this problem. In the paper, we show how to uniquely define the orientation of an arbitrary 2D shape, in terms of what we call its Principal Moments. We show that a small subset of these moments suffice to represent the underlying 2D shape and propose a new method to efficiently compute the shape orientation: Principal Moment Analysis. Finally, we discuss

how this method can further be applied to normalize grey-level images. Besides the theoretical proof of correctness, we describe experiments demonstrating robustness to noise and illustrating the method with real images.

1 Introduction

Representing shape is a challenging task. In fact, unlike local characteristics like color, which can be uniquely determined by a small set of parameters, or texture, which has been successfully captured by using statistical descriptors, the visual information conveyed by a shape, of more global nature and easily perceived by humans, is hard to represent in an appropriate way. This paper deals with two-dimensional (2D) shape representation.

When the 2D shapes to describe are simply connected regions, researchers have used contour-based descriptions (*e.g.*, Chauang and Kuo, 1996; Bartolini et al, 2005). Naturally, for more general shapes, usually consisting in arbitrary sets of points, or landmarks, these descriptors are not adequate. If the points describing the shape are labeled, *i.e.*, if the correspondences between the landmarks of two shapes to compare are known, the problem reduces to the impact of geometrical transformations and disturbances, elegantly addressed through the statistical theory of shape of Kendall et al (1999). However, in many practical scenarios, the shape points are obtained from an automatic process, *e.g.*, edge or corner detection, thus come without

labels or natural ordering.

Estimating the correspondences between points of two shapes leads to a combinatorial problem, which requires prohibitively time-consuming algorithms, even for shapes described by a moderate number of landmarks. To circumvent this problem, researchers have recently attempted to come up with permutation-invariant representations for sets of points. For example, Jebara (2003) shows how to factor out unknown labels through the solution of a convex optimization problem over the set of permutation matrices, and Rodrigues et al (2008a) proposes a permutation-invariant representation obtained by densely sampling an analytic function.

Naturally, the study of shape representation is often motivated by the challenge of comparing two arbitrary shapes. Indeed, if an efficient and universal way of representing shapes is found, the comparison of two shapes can be brought off through the direct comparison of its representations. Regarding universality, an issue that arises when representing and comparing two shapes is their normalization with respect to geometrical transformations, such as translation, scale and rotation. In fact, a desirable requirement of shape representations is that they should form a *complete set of invariants* with respect to these transformations, *i.e.*, two shapes should be equal up to a combination of these transformations *if and only if* their representations are equal. Translation and scale are easily taken care of by removing the mean and normalizing the energy of the shape. Rotation is correspondingly factored out by normalizing according to an *orientation* of the shape, an an-

gle, intrinsic to the shape, that varies coherently when the shape is rotated. However, as we discuss in the following paragraphs, defining and computing the orientation of an arbitrary 2D shape is not so simple.

Additionally, as motivated earlier, complete permutation invariance is a key requirement of unlabeled shape representations as well. Nevertheless, the comparison of two sets of points that are related by an unknown transformation that includes, simultaneously, a 2D rotation, due to different orientation, and a permutation, due to the absence of labels for the points, results highly complex, due to the fact that estimating the transformation leads to a non-convex problem. Iterative methods have been used to compute, in alternate steps, rotation and permutation: the Iterative Closest Point (ICP) algorithm of Besl and McKay (1992), or its probabilistic versions based on Expectation-Maximization (EM) (*e.g.*, McNeill and Vijayakumar, 2006). However, these approaches suffer from the usual sensitivity to the initialization, exhibiting uncertain convergence. Other proposed approaches exhibit drawbacks as well: the convex optimization approach of Jebara (2003) does not deal with rotation and the analytic representation of Rodrigues et al (2008a) is not rotation invariant, requiring pairwise alignment.

The most straightforward method to define orientation uses the principal axis of the shape, obtained, *e.g.*, through Principal Component Analysis (PCA). Although unable to provide a unique orientation, this method defines two possible orientations for elongated shapes, see Fig. 1 (the directional ambiguity happens in general, not only for mirror-symmetric shapes such as

the one used for illustration). For shapes that do not have a well defined principal axis, *e.g.*, rotationally symmetric shapes, PCA-based orientation is completely ambiguous, see Fig. 2.

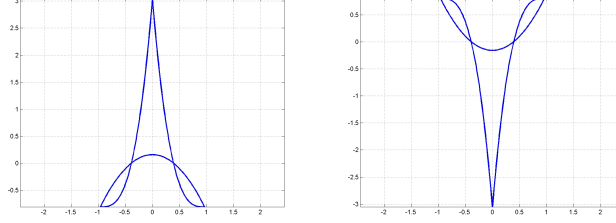


Figure 1: Ambiguity in PCA-based orientation of general shapes: only the principal axis is determined, not its *direction*.

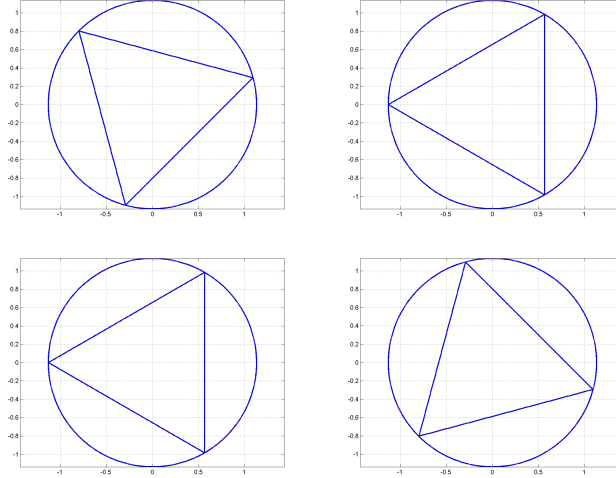


Figure 2: A rotationally symmetric shape. Since for this kind of shapes the *principal axis is not defined*, PCA can not be used to compute their orientation.

To deal with these ambiguities, researchers attempted to work with concepts like mirror-symmetry axes (Atallah, 1985; Marola, 1989), universal

principal axis (Lin, 1993), and generalized principal axis (Tsai and Chou, 1991; Zunic et al, 2006). In general, the motivation for these works is more on the definition of a “reasonable geometric orientation” than on the robust computation of a unique orientation angle for arbitrary shapes. Since rotationally symmetric shapes are particularly challenging, the automatic detection of symmetry and fold number, by itself a relevant problem, has also received attention (Lin et al, 1994; Shen and Ip, 1999; Derrode and Ghorbel, 2004; Prasad and Yegnanarayana, 2004).

The more theoretically sustained methods to compute orientation are based on the geometric moments of the points defining the shape. In particular, the so-called Complex Moments (CMs) were introduced in the eighties (Abu-Mostafa and Psaltis, 1985; Teh and Chin, 1988). The elegance of these approaches comes from defining the orientation through the phase of a single CM of a particular order. In the nineties, more general moments were proposed to deal with degenerate shapes (Shen and Ip, 1997, 1999), at a cost of dealing with several moments, chosen by tuning a free parameter index through search, and detecting rotational symmetry as an intermediate step. However, as we detail in Section 5, these methods do not cope with several shapes that lead to singular moments. In practice, this means that the phase of these moments is sensitive to noise, leading to unstable estimates of orientation. Other approaches require the exhaustive search for the angle maximizing a given orientation measure (Ha and Moura, 2003, 2005), without any guarantee of uniqueness of the solution.

In this paper, we address the need to combine compact descriptions of 2D shapes, for computational efficiency, with the invariance and discriminative power of complete sets of invariants. We propose to represent 2D shapes in terms of particular complex moments, which we call the *Principal Moments* (PMs). Although moments of image patterns have been extensively used due to their invariance properties, since at least the early sixties (Hu, 1962), their discriminative properties have not been studied even in more recent related work (*e.g.*, Abu-Mostafa and Psaltis, 1985; Khotanzad and Hong, 1990; Shen and Ip, 1999; Mukundan et al, 2001). In opposition, our representation uniquely defines the shape. In fact, using the same number of PMs as the number of shape landmarks, our representation forms a complete set of invariants with respect to permutation. We further show that the PMs coincide with the coefficients of the Fourier series of the representation of Rodrigues et al (2008a), a result that guarantees that our representation inherits the discriminative power demonstrated by the experiments reported in that paper. Subsequently, we derive an upper bound for the magnitude of these coefficients in terms of the shape complexity, *i.e.*, of the number of landmarks. Using these results, we show that our representation is compact, in the sense that a small number of PMs (much smaller than the number of landmarks) suffices to represent the shape. This compactness contrasts with the usual large dimension of other complete representations (*e.g.*, those based on the bispectrum (Kondor, 2008) or the densely sampled functions of Rodrigues et al (2008a)), an issue of outmost relevance when working with

large databases.

In what respects to geometric transformations, besides trivially extending the representation to deal with translation and scale transformations, a major contribution of this paper is the extension to also include rotations. This extension consists in previously rotating each shape instance of a normalization angle which defines the shape orientation. Overcoming the limitations discussed above, we present a new method to define and compute a unique orientation of any 2D shape, based on its PMs. More specifically, we show that the phases of two of these moments unambiguously define the orientation of an arbitrary 2D shape (including rotationally symmetric ones) and propose an algorithm, *Principal Moment Analysis* (PMA), that computes the orientation angle by integrating the contributions of all pairs of moments.

Naturally, PMA can be used to normalize arbitrary 2D shapes, *e.g.*, binary images, with respect to orientation, before any other processing takes place. In the paper we also discuss the straightforward extension of PMA to the normalization of general gray-level images.

Besides theoretically sound, PMA results are robust to noise, as the experiments in the paper illustrate.

The remaining of the paper is organized as follows. Section 2 introduces the PMs and relates them to previously proposed moments. In Section 3, we relate the PMs with the representation of Rodrigues et al (2008a) and derive an upper bound for the length containing most of its energy, which enables ending up with a compact shape representation. Section 4 describes our

approach to extend the representation towards obtaining a complete set of invariants with respect to 2D rotation. In Section 5, we detail the limitations of current methods when estimating 2D orientation. Section 6 presents PMA, our algorithm for computing a unique orientation of an arbitrary 2D shape from its PMs. In Section 7, we extend PMA to the orientation normalization of gray-level images. Section 8 contains experiments and Section 9 concludes the paper.

2 Principal Moments for 2D Shape Representation

Consider an arbitrary 2D shape described by a set of N points in the plain, thus by an N -dimensional complex vector $\mathbf{z} \in \mathbb{C}^N$, containing their coordinates:

$$\mathbf{z} = \begin{bmatrix} x_1 + jy_1 \\ x_2 + jy_2 \\ \vdots \\ x_N + jy_N \end{bmatrix} = \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_N \end{bmatrix}. \quad (1)$$

Naturally, since the points do not have labels, the same shape can be described by any vector obtained from \mathbf{z} by re-ordering its entries.

We define the k^{th} -order *Principal Moment* (PM), $k \in \{1, 2, 3, \dots\}$, by

$$M_k(\mathbf{z}) = \frac{1}{Nk!}(z_1^k + z_2^k + \dots + z_N^k) = \frac{1}{Nk!} \sum_{n=1}^N z_n^k. \quad (2)$$

The k^{th} PM, after stripping off the scaling factor $1/(Nk!)$, is also known as the k^{th} *power sum* of the landmarks $\{z_1, z_2, \dots, z_N\}$. The inclusion of this particular scaling factor in the definition (2) is motivated in the next section. Through the Newton's identities, the first N power sums can be unambiguously converted into the so-called *fundamental symmetric polynomials* of the landmarks. As these polynomials can be seen as the coefficients (up to sign changes) of the N^{th} -order univariate polynomial containing the landmarks as roots, there is a bijection between this univariate polynomial and the first N power sums. The reader is referred to the enlightening book of Kanatani (1990) for what respects to these equivalences. Since an N^{th} -order univariate polynomial unequivocally defines its N roots, although orderless, the first N power sums unambiguously define the shape \mathbf{z} , up to a permutation, and so do the first N PMs. This N -sized representation is thus said to form a *complete set of invariants* over the permutation group, or, equivalently, to be *maximally invariant* to permutations. In other words, (i) any vector obtained from \mathbf{z} by re-ordering its entries leads to the same PMs (permutation invariance); and (ii) any vector with at least one landmark lying at a different position than the ones in \mathbf{z} , leads to distinct PMs (discrimination).

The defined representation can be trivially extended to be maximally

invariant with respect to simple geometric transformations. Indeed, it is easily shown that the simple pre-processing step of working with

$$\sqrt{N} \frac{\mathbf{z} - \bar{\mathbf{z}}}{\|\mathbf{z} - \bar{\mathbf{z}}\|}, \quad (3)$$

where $\bar{\mathbf{z}} = \frac{1}{N} \sum_{n=1}^N z_n$ denotes the center of mass of \mathbf{z} and $\|\cdot\|$ denotes the l_2 norm, rather than directly with \mathbf{z} , extends the complete set of invariants to include translation and scale transformations. The extension with respect to rotation operations, though, is not so trivial and it is, along with the subject of shape representation using PMs, the central subject of this paper.

Before proceeding, we relate the PMs in (2) with the more general *Complex Moments* (CMs) of Abu-Mostafa and Psaltis (1985) and *Generalized Complex* (GC) moments of Shen and Ip (1997, 1999). The CM of order (p, q) of an image $g(x, y)$ is defined by

$$C_{pq}(g) = \int \int_{-\infty}^{+\infty} (x + jy)^p (x - jy)^q g(x, y) dx dy, \quad (4)$$

where $p \geq 0$ and $0 \leq q \leq p$ (Abu-Mostafa and Psaltis, 1985). Considering an image composed by a set of N mass points located at the shape landmarks $\{z_1, z_2, \dots, z_N\}$, the integral in (4) becomes a sum:

$$C_{pq}(\mathbf{z}) = \sum_{n=1}^N z_n^p (z_n^*)^q, \quad (5)$$

where z_n^* denotes the complex conjugate of z_n . In turn, the GC moment of

order (p, q) is given by the integral in polar coordinates

$$GC_{pq}(g) = \int_{-\pi}^{\pi} \int_0^{\infty} r^p e^{jq\theta} g(r \cos \theta, r \sin \theta) r dr d\theta, \quad (6)$$

where $p \in \{0, 1, 2, \dots\}$ and $q \in \{1, 2, 3, \dots\}$ (Shen and Ip, 1997). For a shape \mathbf{z} , the GC moments collapse into the sums

$$GC_{pq}(\mathbf{z}) = \sum_{n=1}^N |z_n|^p e^{jq \arg z_n}. \quad (7)$$

It is now clear that the PMs are CMs and GC moments of particular orders (up to a scaling factor): from (2,5,7), $M_k(\mathbf{z}) = C_{k0}(\mathbf{z})/(Nk!) = GC_{kk}(\mathbf{z})/(Nk!)$. Finally, note also that, for any shape, $M_0(\mathbf{z}) = \frac{1}{N} \sum_{i=1}^N 1 = 1$, and, assuming the shapes were pre-processed as in (3), $M_1(\mathbf{z}) = \frac{1}{N} \sum_{n=1}^N z_n = 0$.

3 A Compact Shape Representation

We now show why a representation that uses a small subset of PMs suffices in practice. To do this, we start by relating the PMs with the *analytic signature* (ANSIG), the representation introduced by Rodrigues et al (2008a). The ANSIG is an analytic function on the complex plane, obtained from \mathbf{z} through

$$a(\mathbf{z}, \xi) = \frac{1}{N} \sum_{n=1}^N e^{z_n \xi}. \quad (8)$$

Rodrigues et al (2008a) show that this representation is complete and permutation-invariant and thoughtfully illustrate its capabilities with several shape-based image classification experiments. In these experiments, the analytic function (8) is described in the computer by its 512 samples uniformly taken on the unit-circle of the complex plane. Working with these high dimensional vectors may be adequate for tasks requiring the comparison of a small number of shapes but certainly not for applications that deal with very large databases, *e.g.*, the internet.

In the sequel, we derive that our PMs (2) are intimately related with the ANSIG (8) and show that a small number of either PMs or of ANSIG samples suffices to represent the shape with similar performance.

3.1 The Principal Moments and the ANSIG Spectrum

A direct consequence of Cauchy’s integral formula (see, *e.g.*, Ahlfors, 1978) is that any analytic function is fully specified by the values it takes on a closed contour on the complex plane. Thus, the ANSIG in (8) is equivalently described by its restriction to the unit-circle,

$$h(\mathbf{z}, \theta) = a(\mathbf{z}, e^{j\theta}) = \frac{1}{N} \sum_{n=1}^N \exp(z_n e^{j\theta}) . \quad (9)$$

To approximate the unit-circle restriction of the ANSIG by a finite-dimensional computer representation, we study its frequency spectrum. Note that $h(\mathbf{z}, \theta)$ in (9) can be seen as a real-argument complex-valued peri-

odic function, with fundamental period $T = 2\pi$ and fundamental frequency $\omega_0 = 2\pi/T = 1$. Thus, it can be written in terms of its Fourier series,

$$h(\mathbf{z}, \theta) = \sum_{k=-\infty}^{+\infty} H_k(\mathbf{z}) e^{jk\theta}, \quad (10)$$

where each coefficient H_k is given by (see, *e.g.*, Oppenheim et al, 1996):

$$H_k(\mathbf{z}) = \frac{1}{2\pi} \int_{-\pi}^{\pi} h(\mathbf{z}, \theta) e^{-jk\theta} d\theta. \quad (11)$$

The analysis expression (11) is hard to carry out in the case of $h(\mathbf{z}, \theta)$ given by (9). However, the coefficients of the Fourier series easily follow from the comparison of the synthesis expression (10) with the definition of $h(\mathbf{z}, \theta)$ in (9), after some manipulations. In fact, expressing the exponential in (9) by its Maclaurin series (*i.e.*, its Taylor series at the origin), we get

$$\begin{aligned} h(\mathbf{z}, \theta) &= \frac{1}{N} \sum_{n=1}^N \sum_{k=0}^{\infty} \frac{(z_n e^{j\theta})^k}{k!} \\ &= \sum_{k=0}^{\infty} \left(\frac{1}{N k!} \sum_{n=1}^N z_n^k \right) e^{jk\theta}. \end{aligned} \quad (12)$$

Comparing (12) with (10), and since the coefficients of the Fourier series representation are unique (Oppenheim et al, 1996), we conclude that

$$H_k(\mathbf{z}) = \begin{cases} \frac{1}{N k!} \sum_{n=1}^N z_n^k & \text{for } 0 \leq k < +\infty, \\ 0 & \text{for } -\infty < k \leq -1. \end{cases} \quad (13)$$

Thus, given a 2D shape described by a set of N landmarks, expression (13) relates the values of the coefficients of the Fourier series of the unit-circle restriction of its ANSIG, $\{H_k, -\infty < k < +\infty\}$, to the positions of the landmarks in the plane, $\{z_1, z_2, \dots, z_N\}$. Comparing with (2), we conclude that these coefficients for $k \geq 0$ coincide with the PMs:

$$H_k(\mathbf{z}) = M_k(\mathbf{z}), \quad k = 0, 1, 2, \dots$$

Due to the one-to-one correspondence just proved between $M_k(\mathbf{z})$, $h(\mathbf{z}, \theta)$ and $a(\mathbf{z}, \xi)$, the PM shape representation inherits all the properties of the ANSIG. Although the fact that the PMs constitute a complete representation was already referred in the previous section (and the fact that they enjoy permutation invariance is immediate from definition (2)), we now know they are equivalent to the ANSIG also in what respects to discrimination capabilities.

3.2 Compactness of the Principal Moments

We now show that the proposed representation is compact, in the sense that only a small number of PMs is enough to represent the shape with neglectable loss of discrimination power. For that purpose, we derive an upper bound $b(k)$ for the magnitude of the coefficient H_k (or, equivalently, the PM M_k), for $k \geq 2$, in terms of the number of landmarks describing the shape. Using this result, we deliver an upper limit for the (wide-sense) bandwidth of the unit-circle restriction $h(\mathbf{z}, \theta)$ of the ANSIG. The first step is done through

the following chain of equalities and inequalities:

$$|M_k| = \frac{1}{Nk!} \left| \sum_{n=1}^N z_n^k \right| \leq \frac{1}{Nk!} \sum_{n=1}^N |z_n|^k \quad (14)$$

$$= \frac{1}{Nk!} \sum_{n=1}^N (|z_n|^2)^{\frac{k}{2}} \leq \frac{1}{Nk!} \left(\sum_{n=1}^N |z_n|^2 \right)^{\frac{k}{2}} \quad (15)$$

$$= \frac{N^{\frac{k}{2}-1}}{k!} \stackrel{\text{def}}{=} b(k), \quad (16)$$

where (14) uses the triangle inequality, (15) uses the fact that the function $(a+b)^n$ is convex for a, b real positive and $n \geq 1$, and (16) is due to the fact that $\sum_{n=1}^N |z_n|^2 = N$, for shape vectors \mathbf{z} normalized according to (3).

To estimate the bandwidth of $h(\mathbf{z}, \theta)$ in terms of the number of landmarks, usually, we would seek the smallest k such that the ratio $|M_k|/|M_0|$ is below a given threshold p . In our case, since $M_0 = 1$, this would lead to $|M_k| < p$. Nevertheless, as we do not know the exact value of $|M_k|$, we will use the upper bound $b(k)$ as a proxy. We are not sure to obtain the smallest k that satisfies our bandwidth constraint, but we guarantee the satisfaction of the inequality, since $|M_k| \leq b(k)$.

Finding the smallest k such that $b(k) < p$ requires solving the limit case equation $N^{\frac{k}{2}-1}/k! = p$, for which there is no analytic solution. We propose a simple method to solve for k numerically. Due to the fast grow of $k!$ and to increase stability, we apply logarithms on both sides of the inequality.

Denoting the natural logarithm of $b(k)$ by $B(k)$, we have:

$$B(k) = \ln b(k) = \left(\frac{k}{2} - 1\right) \ln N - \ln k! \quad (17)$$

$$\begin{aligned} &\simeq k \left(\frac{\ln N}{2} + 1\right) - k \ln k \\ &\quad - \frac{1}{2} \ln k - \ln N - \frac{1}{2} \ln 2\pi, \end{aligned} \quad (18)$$

where (17) uses the definition of $b(k)$ in (16) and (18) uses the Stirling's approximation $\ln k! \simeq k \ln k - k + \frac{1}{2} \ln 2\pi k$ (see, *e.g.*, Paris and Kaminsky, 2001).

To analyze the behavior of $B(k)$ given by (18), we relax k to the reals and express the first two derivatives:

$$B'(k) = \frac{\ln N}{2} - \ln k - \frac{1}{2k}, \quad B''(k) = -\frac{1}{k} + \frac{1}{2k^2}.$$

From these expressions, we see that $B(k)$ has an inflection at $k = 1/2$, where $B''(k) = 0$, and two extrema at $k = k_1 < 1/2$ and $k = k_2 > 1/2$, where $B'(k) = 0$ (assuming $N \geq 2$). Furthermore, $B(k)$ monotonically decreases for $k > k_2$, where $B'(k) < 0$, being $\lim_{k \rightarrow +\infty} B(k) = -\infty$. The plot in Fig. 3 illustrates the behavior of $B(k)$ for $N = 10$. To find k such that $B(k)$ is below a given threshold $\ln p$, we thus propose the following strategy in two steps: first, solve $B'(k) = 0$ in the interval $k \in [1/2, +\infty)$, obtaining k_2 . Then, solve $B(k) - \ln(p) = 0$ in the interval $k \in [k_2, +\infty)$, obtaining $k = k_B$, the desired upper bound for the bandwidth of the ANSIG. Note that the first

step is necessary to specify the lower limit of the search region for the second one: without that limit, we could obtain a spurious solution $k_B < k_2$.

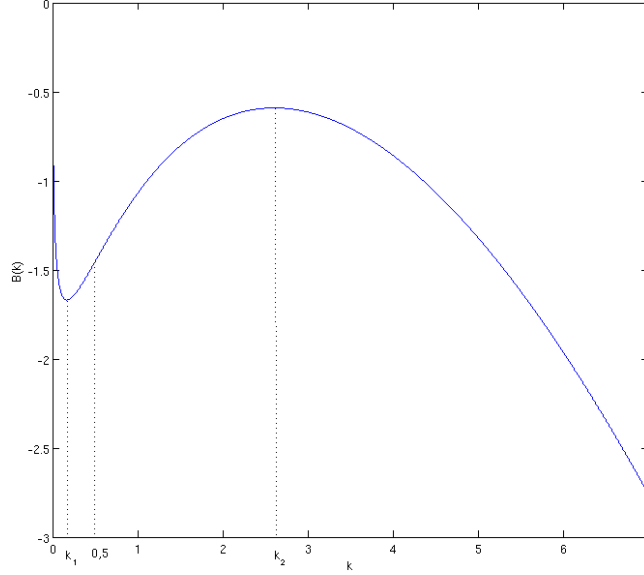


Figure 3: Upper bound $B(k)$ for the (logarithm of the) magnitude of the spectrum of the analytic signature (or, equivalently, magnitude of the principal moments), for $N = 10$ points.

We thus conclude that most of the energy (the parameter p controls the amount) of the ANSIG of a shape described by N landmarks is contained in a number k_B of complex coefficients, which, naturally, depends on N . Fig. 4 plots the number of coefficients k_B , computed as described above, as a function of the number of landmarks N , for $p = 0.1$ (-20dB). Obviously, k_B can be indistinctly interpreted as either the required number of Fourier series coefficients, *i.e.*, the number of PMs, or the required number of samples in the unit-circle to represent the shape. In fact, since the fundamental frequency is $\omega_0 = 1$, the approximate bandwidth of the signal is $\omega_B = k_B \omega_0 = k_B$.

Since the spectrum of $h(\mathbf{z}, \theta)$ is zero for negative frequencies, *c.f.* (13), it suffices to sample at a rate (number of points) of $N = \omega_s = \omega_B = k_B$ (the Nyquist sampling rate ω_s of twice the bandwidth is only required for two-sided spectra (Oppenheim et al, 1996, 1999)). Naturally, to recover the original continuous ANSIG from these samples, we should use a (complex coefficient) filter with passband $\omega \in [0, \omega_s)$ (in opposition to the traditional low-pass filter with cutting frequency $\omega_s/2$).

The plot in Fig. 4 also compares the required number k_B of samples, or of PMs, with 512, the fixed number of samples used in Rodrigues et al (2008a): while for shapes described by an huge number of points (more than $\simeq 40000$), 512 samples may not be enough, for the majority of cases that may arise in practice (a few hundreds of landmarks), the required number is much smaller (a few dozens). Note further that k_B is smaller than N , making the representation based on PMs loose its maximal invariance to permutations, in a strict sense, when using k_B coefficients. Nevertheless, the discrimination loss that results from this is small, since most of the energy of the signature is captured by the first k_B coefficients ¹.

Having motivated the usage of only k_B PMs as a way to compactly represent shapes in terms of its complexity, *i.e.*, its number of landmarks, we now discuss the comparison of descriptions of shapes of distinct complexity. Since the first k_B PMs are (the most relevant) coefficients of the ANSIG

¹Naturally, the compactness of the representation is due to the decay of $|M_k|$ with $k!$ imposed by the normalization factor in (2), which is now motivated. In Appendix A, we further discuss the issue of normalizing power sums.

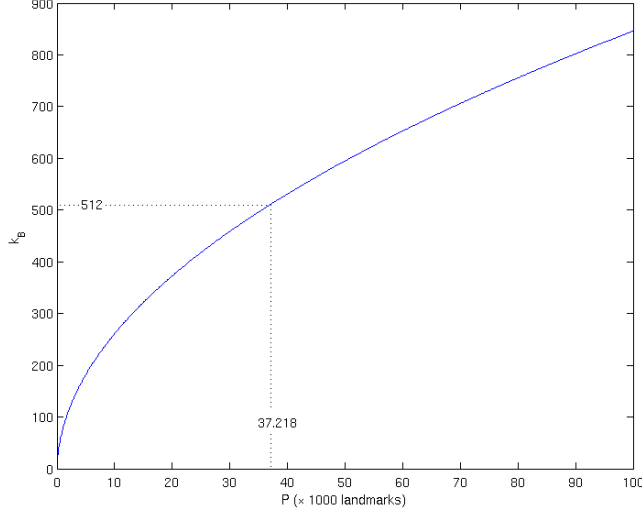


Figure 4: The number of coefficients k_B needed to represent a shape described by N landmarks.

Fourier series, it suffices to pad with zeros the smaller vector before performing the comparison in frequency domain. If, in opposition, the shapes are equivalently described by the sparse set of k_B ANSIG samples, say N_1 samples for one of the shapes and N_2 samples for the other, it is necessary to use multirate signal processing techniques to convert both to a common sampling rate (see, *e.g.*, Oppenheim et al, 1999). For example, perform up-sampling of the ANSIGs by a factor of, respectively, $L_1 = \text{lcm}(N_1, N_2)/N_1$ and $L_2 = \text{lcm}(N_1, N_2)/N_2$, where lcm stands for the least common multiple, followed by interpolation using (complex coefficient) filters with passband, respectively, $\omega \in [0, 2\pi/L_1)$ and $\omega \in [0, 2\pi/L_2)$.

4 Maximal Rotation Invariance

The representation of a 2D shape by its PMs forms a complete set of invariants with respect to permutation. The extension of the set to include translation and scale invariance is easily obtained by removing the mean and normalizing the energy, according to (3). However, taking care of the orientation is not obvious. In the following, we propose an extension to include maximal invariance to rotations by computing a unique orientation, definable for an arbitrary shape.

A natural and the most common way to attempt to obtain rotation invariance consists in finding an angle $\theta(\mathbf{z})$ such that, through rotation, any shape \mathbf{z} is brought to its “normalized” version

$$\mathbf{w}(\mathbf{z}) = \mathbf{z}e^{-j\theta(\mathbf{z})}. \quad (19)$$

In fact, if the desired invariance is satisfied, *i.e.*, if,

$$\forall \phi, \quad \mathbf{w}(\mathbf{z}e^{j\phi}) = \mathbf{w}(\mathbf{z}), \quad (20)$$

the normalization in (19) produces a maximal invariant. However, current methods to compute the shape orientation $\theta(\mathbf{z})$ either fail to process particular shapes (see examples in Figs. 1 and 2 and others in the following section) or do not guarantee the equality in (20).

The success of the normalization in (19) hinges then on finding an ap-

appropriate function $\theta : \mathbb{C}^N \rightarrow (-\pi, \pi]$ that unambiguously defines $\theta(\mathbf{z})$, the orientation of the shape \mathbf{z} . We now show that, to guarantee the desired invariance, it suffices that this function satisfies a natural condition: that the orientation of a rotated shape is the sum of the orientation of the original shape with the rotation angle. Formally, this condition means that the function $\theta(\cdot)$ has to satisfy

$$\forall \phi, \quad \theta(\mathbf{z}e^{j\phi}) = \theta(\mathbf{z}) + \phi, \quad (21)$$

where the equality is modulo 2π . Simple manipulations show that (21) suffices to guarantee (20):

$$\mathbf{w}(\mathbf{z}e^{j\phi}) = \mathbf{z}e^{j\phi} \exp(-j\theta(\mathbf{z}e^{j\phi})) \quad (22)$$

$$= \mathbf{z}e^{j\phi} e^{-j(\theta(\mathbf{z})+\phi)} \quad (23)$$

$$= \mathbf{z}e^{-j\theta(\mathbf{z})}$$

$$= \mathbf{w}(\mathbf{z}), \quad (24)$$

where (22) and (24) use the definition (19) and (23) uses (21).

In Section 6, we propose an orientation function $\theta(\cdot)$ satisfying (21). This is a relevant contribution to maximally invariant representations for 2D shapes, since the pre-processing step (19) extends any permutation-invariant representation, like the ones proposed by Jebara (2003) and Rodrigues et al (2008a), or the PMs introduced in Section 2 of this article, to also accommo-

date maximal invariance with respect to arbitrary geometric transformations, *i.e.*, including rotation.

5 Limitations of Previous Approaches to Moment-based Orientation Normalization

To obtain the orientation $\theta(\mathbf{z})$ of a shape \mathbf{z} , we use the PMs of the points describing the shape. Since image moments have been used in the past, we first overview moment-based estimation of orientation and motivate the need to revisit the problem. The usage of Complex Moments (CMs) to define orientation was proposed in Abu-Mostafa and Psaltis (1985). CMs stand for compact representations of linear combinations of ordinary (*i.e.*, real) geometric moments. In that work, the authors define and compute the orientation by imposing the phase of one of the moments $C_{q+1,q}$ in (4) to be zero. When applying this method to a shape \mathbf{z} , *i.e.*, to an image composed by a set of N mass points describing the shape, we obtain, through (5), the moments

$$C_{q+1,q}(\mathbf{z}) = \sum_{n=1}^N |z_n|^{2q+1} e^{j \arg z_n}, \quad (25)$$

where, as introduced in Section 2, z_n collects the coordinates of the n^{th} shape point.

Although the method just described is adequate to deal with shapes \mathbf{z} that lead to a moment $C_{q+1,q}(\mathbf{z})$ with large magnitude, there are shapes for

which this does not happen for any q . It was known that this is the case of rotationally symmetric shapes (Abu-Mostafa and Psaltis, 1985), but we now show it may also happen with general ones. Just look at the example in Fig. 5, where $\mathbf{z} = [1, j, -j, \exp(j2\pi/3), \exp(-j2\pi/3)]^T$. For this shape, from (25), we obtain that all the moments $C_{q+1,q}(\mathbf{z})$ are zero, regardless of q :

$$C_{q+1,q}(\mathbf{z}) = e^{j0} + e^{j\pi/2} + e^{-j\pi/2} + e^{j2\pi/3} + e^{-j2\pi/3} = 0.$$

For S -fold rotationally symmetric shapes, Abu-Mostafa and Psaltis (1985) propose to use the phase of one of the moments $C_{q+S,q}$. However, again, there are S -fold rotationally symmetric shapes for which these are all zero. For example, it is straightforward to show that the 2-fold rotationally symmetric shape in Fig. 6 leads to $C_{q+2,q}(\mathbf{z}) = 0, \forall q$. Although the examples in Figs. 5 and 6 (or others the reader may come up with) serve as mere illustrations of extreme cases, they also make clear that in practice it is not adequate to rely on the angle of these moments to robustly compute shape orientation, since when the magnitude of those moments is small, their phase results very sensitive to the noise.

Moment-based orientation was later addressed by using Generalized Complex (GC) moments (Shen and Ip, 1997, 1999). GC moments, simply termed rotational moments in a previous review (including Legendre, Zernike, and CMs) by Teh and Chin (1988) are given by (6) and can be seen as the coefficients of the Fourier series of radial projections of the image. To deal

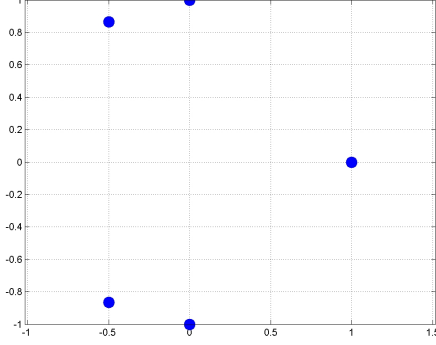


Figure 5: Example of a shape (composed by the points $\{1, \pm j, \exp(\pm j2\pi/3)\}$), for which all the complex moments $C_{q+1,q}$ are zero, making impossible to define its orientation in terms of the phase of such moments.

with ambiguities that arise when attempting to define and compute shape orientation from a single moment of a particular order, Shen and Ip (1997, 1999) use three non-zero GC moments with a fixed index p . The method is not simple: from GC_{pq_1} and GC_{pq_2} , *the possibility* is inferred that the shape is rotationally symmetric; in case there is that possibility, the unambiguous detection of symmetry requires an exhaustive search; if the shape is classified as rotationally symmetric, a third moment GC_{pq_3} is also used to compute the orientation. The simple example in Fig.7 shows that this method may fail: consider $\mathbf{z}_1 = [1, -1/4, -3/4]^T$ and \mathbf{z}_2 its reflection, *i.e.*, \mathbf{z}_1 rotated by π , $\mathbf{z}_2 = -\mathbf{z}_1$, and the choice of GC index $p = 1$. Using (7), we get

$$GC_{1q}(\mathbf{z}_1) = 1 + \frac{1}{4}e^{jq\pi} + \frac{3}{4}e^{jq\pi} = 1 + (-1)^q = GC_{1q}(\mathbf{z}_2),$$

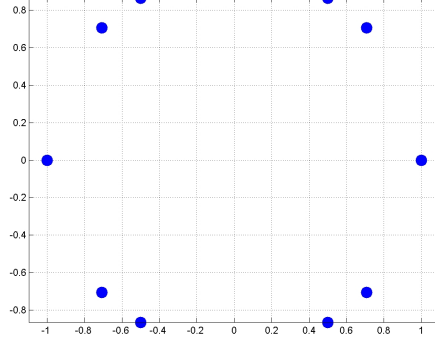


Figure 6: Example of a 2-fold rotationally symmetric shape (composed by the points $\{\pm 1, \pm \exp(\pm j\pi/3), \pm \exp(\pm j\pi/4)\}$), for which all the complex moments $C_{q+2,q}$ are zero, making impossible to define its orientation in terms of the phase of such moments.

showing that it is impossible to distinguish between the orientations of \mathbf{z}_1 and \mathbf{z}_2 from the moments GC_{1q} (note that the shape in \mathbf{z}_1 and \mathbf{z}_2 is not rotationally symmetric, thus different orientation angles $\theta(\mathbf{z}_1)$ and $\theta(\mathbf{z}_2)$ must be computed). Although Shen and Ip (1997, 1999) propose to tune the index p by maximizing a so-called alternating energy of the radial projection (which also requires exhaustive search), this method fails to exclude $p = 1$ for the example above.

6 The Unique Orientation of a 2D Shape Through Principal Moment Analysis

We now present our algorithm to compute the unique orientation of an arbitrary shape, *i.e.*, we derive a function $\theta(\mathbf{z})$ that satisfies property (21). Our

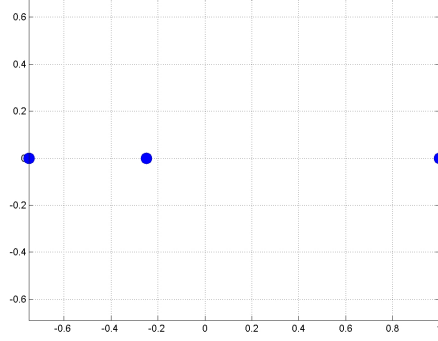


Figure 7: Example of a shape (composed by the points $\{1, -1/4, -3/4\}$), for which all the generalized complex moments GC_{1q} are the same as for its reflexion ($\{-1, 1/4, 3/4\}$), illustrating that it is impossible to unambiguously define their orientations in terms of those moments.

approach to define and compute the orientation $\theta(\mathbf{z})$ is based on the PMs of the shape, given by (2). As we will only deal with their argument, for simplicity, we strip off the scaling factor and work directly with the power sums

$$\mu_k(\mathbf{z}) = z_1^k + z_2^k + \cdots + z_N^k = \sum_{n=1}^N z_n^k, \quad (26)$$

where $k \in \{1, 2, 3, \dots\}$. Note, nevertheless, that both the power sums $\{\mu_k(\mathbf{z})\}$ (26) and the PMs $\{M_k(\mathbf{z})\}$ (2) lead to the same algorithm to be explained, and that both forms can be used to compute the (same) unique orientation of the shape. For this reason, in this section, we refer to $\mu_k(\mathbf{z})$ as the k^{th} -order PM of the shape.

6.1 A Single Moment is Not Sufficient

If we were to choose $\theta(\mathbf{z}) = \arg \mu_1(\mathbf{z})$ (equivalent to applying the simplest form of the method in Abu-Mostafa and Psaltis (1985), using $C_{10} = \mu_1$), we would satisfy (21). In fact, from (26), $\mu_1(\mathbf{z}e^{j\phi}) = \mu_1(\mathbf{z})e^{j\phi}$, thus $\arg \mu_1(\mathbf{z}e^{j\phi}) = \arg \mu_1(\mathbf{z}) + \phi$. Nevertheless, since $\mu_1(\mathbf{z}) = \sum_n z_n$ is proportional to the shape center, its angle is not a characteristic of the shape format, but only of the shape localization. In practice, to obtain translation invariance, it is common to apply the pre-processing step (3), where the shapes are centered by subtracting their center of mass, thus μ_1 is zero for all shapes and useless to determine a shape orientation.

The choice $\theta(\mathbf{z}) = \arg \mu_1(\mathbf{z})$ is equivalent to imposing the argument of the first-order PM of the rotationally normalized shape (19) to be zero, *i.e.*, imposing $\arg \mu_1(\mathbf{z}e^{-j\theta(\mathbf{z})}) = 0$. Our approach is to generalize this method by doing the same to the k^{th} -order PM, assumed to be non-zero:

$$\arg \mu_k(\mathbf{z}e^{-j\theta(\mathbf{z})}) = 0. \quad (27)$$

Note that one non-zero PM must exist, as otherwise all landmarks of the shape are at the origin. To solve for $\theta(\mathbf{z})$, use the definition (26) to rewrite (27) as

$$\arg \sum_{n=1}^N z_n^k e^{-jk\theta(\mathbf{z})} = 0.$$

Since complex arguments are defined modulo 2π , we get

$$\arg \sum_{n=1}^N z_n^k - k\theta(\mathbf{z}) + 2\pi l = 0, \quad (28)$$

where l is an integer. Now, we express the solution(s) for the normalization angle as

$$\theta(\mathbf{z}) = \frac{\arg \mu_k(\mathbf{z})}{k} + \frac{2\pi}{k}l, \quad l \in \{0, 1, \dots, k-1\}, \quad (29)$$

where we used (26) again and noted that only k values of l lead to distinct solutions for $\theta(\mathbf{z})$. Expression (29) makes clear that an ambiguity arises when attempting to define the normalization angle using the k^{th} -order PM alone (for $k \neq 1$): there are k (modulo 2π) different values of $\theta(\mathbf{z})$ that annihilate the argument of $\mu_k(\mathbf{z}e^{-j\theta(\mathbf{z})})$.

As previously referred, the normalization angle $\theta(\mathbf{z})$ needs to satisfy property (21), *i.e.*, the normalization angle of a rotated shape must be equal to the one of the original shape plus the rotation angle. We derive what this condition imposes to the solution for $\theta(\mathbf{z})$ that must be picked from the set in (29). By proceeding in a similar way as in (27)–(28), we express the argument of the k^{th} -order PM of a rotated shape as

$$\arg \mu_k(\mathbf{z}e^{j\phi}) = \arg \mu_k(\mathbf{z}) + k\phi + 2\pi\hat{l}, \quad (30)$$

where \hat{l} is an integer that guarantees that the argument of $\mu_k(\mathbf{z}e^{j\phi})$ falls within the interval where this operator is defined, *e.g.*, $(-\pi, \pi]$. Using this

result and (29), we obtain the normalization angle of the rotated shape:

$$\theta(\mathbf{z}e^{j\phi}) = \frac{\arg \mu_k(\mathbf{z})}{k} + \phi + \frac{2\pi}{k}(l + \hat{l}). \quad (31)$$

We now see that the verification of the desired property (21) hinges on the choice of the integer l in the definition (29) of the normalization angle. In fact, if one decides to simply pick a given fixed l , *i.e.*, to choose the same l for all shape vectors \mathbf{z} , (31) becomes $\theta(\mathbf{z}e^{j\phi}) = \theta(\mathbf{z}) + \phi + (2\pi/k)\hat{l}$, showing that property (21) is satisfied if and only if $\hat{l} = 0 \pmod{k}$. However, this can not be guaranteed, since in general equality (30) requires distinct values of \hat{l} for distinct ϕ : just imagine ϕ ranging from 0 to 2π and note that $\arg \mu_k(\mathbf{z}e^{j\phi})$ would exhibit jumps (in order to maintain its value within $(-\pi, \pi]$) corresponding to a changing value of \hat{l} , at values of ϕ spaced by intervals of length $2\pi/k$.

6.2 Using a Pair of Moments

The crux of our approach is to define the normalization angle in (29) by selecting a value for l that *depends on the shape* vector \mathbf{z} . We will show that it is always possible to select $l(\mathbf{z})$ according to our method, *i.e.*, that it unambiguously defines a normalization procedure, and that the resulting normalization angle $\theta(\mathbf{z})$ satisfies property (21), thus guaranteeing maximal invariance to rotations in shape-based classification.

To achieve this, we use a supplementary non-zero PM, $\mu_m(\mathbf{z})$, with k and

m coprime, *i.e.*, with $\gcd(k, m) = 1$, where \gcd denotes the greatest common divisor. The case where there do not exist coprime non-zero moments will be treated in the next subsection. Note additionally that there are no shapes (with a finite number of points) with all but one non-zero PM². Consequently, all shapes (except the one with all landmarks at the origin) have at least two non-zero moments, although not compulsorily coprime.

Our choice for $l(\mathbf{z})$ is based on the arguments of the k^{th} - and m^{th} -order PMs. For simplicity, denote the argument of the m^{th} -order PM of the normalized shape by $\nu(\mathbf{z}, l)$, where $l \in \{0, 1, \dots, k-1\}$ stands for the particular integer used in (29) to compute the normalization angle. Then,

$$\begin{aligned} \nu(\mathbf{z}, l) &= \arg \mu_m(\mathbf{z} e^{-j\theta(\mathbf{z}, l)}) \\ &= \arg \sum_{n=1}^N z_n^m - m\theta(\mathbf{z}, l) \end{aligned} \quad (32)$$

$$= \arg \mu_m(\mathbf{z}) - \frac{m}{k} \arg \mu_k(\mathbf{z}) - \frac{m}{k} 2\pi l, \quad (33)$$

where we used the definitions of the PMs (26) in (32) and of the normalization angle (29) in (33), emphasizing its dependence on l . It is easy to verify that, since k and m are coprime, the set $\{-(m/k)2\pi l : l = 0, 1, \dots, k-1\}$ is the same as the one expressed by $\{(1/k)2\pi l : l = 0, 1, \dots, k-1\}$ modulo 2π ,

²A simple way to show this is to use the fact that the first N PMs fully specify a shape with N points (Kanatani, 1990). Consider all of them are 0 but $\mu_k = 1$, for an arbitrary choice of k . From this, compute the shape and from the shape compute the remaining (higher order) PMs (or compute these directly using Newton's identities). It will be clear from the resulting expression that they can not be all 0.

i.e., for each element of the first set, there is an element of the second one that differs by a multiple of 2π and vice-versa. Thus, the set $\{\nu(\mathbf{z}, l) : l = 0, 1, \dots, k-1\}$ contains k different elements (mod 2π) spaced by intervals of length $2\pi/k$. We propose to unambiguously choose $l(\mathbf{z})$ so that $\nu(\mathbf{z}, l(\mathbf{z}))$ falls within an arbitrary but *fixed* interval of length $2\pi/k$:

$$\nu(\mathbf{z}, l(\mathbf{z})) \in I \pmod{2\pi}, \quad (34)$$

where the interval

$$I = \{\lambda : \lambda_0 < \lambda \leq \lambda_0 + 2\pi/k\},$$

defined by an arbitrary but fixed $\lambda_0 \in \mathbb{R}$, is independent of \mathbf{z} .

The ambiguity in the definition of the normalization angle $\theta(\mathbf{z})$ in (29) is now solved through the choice of $l(\mathbf{z})$ in (34), but we still have to check that this solution satisfies property (21), *i.e.*, that the normalization angle of a rotated shape equals the one of the original shape plus the rotation angle. As derived above, the normalization angle of a rotated shape is given by

$$\theta(\mathbf{z}e^{j\phi}) = \frac{\arg \mu_k(\mathbf{z})}{k} + \phi + \frac{2\pi}{k} \left(l(\mathbf{z}e^{j\phi}) + \hat{l} \right). \quad (35)$$

This last expression is a simple rewrite of (31), now emphasizing the dependence of l on the (rotated) shape; $l(\mathbf{z}e^{j\phi})$ is thus the solution of (34) for the rotated shape $\mathbf{z}e^{j\phi}$. To express the right side of (35) in terms of $\theta(\mathbf{z})$, we

must relate $l(\mathbf{z}e^{j\phi})$ to $l(\mathbf{z})$. This is done by expressing the argument of the m^{th} PM of the normalized rotated shape in terms of the one of the original shape:

$$\begin{aligned} \nu(\mathbf{z}e^{j\phi}, l(\mathbf{z}e^{j\phi})) &= \\ &= \arg \mu_m(\mathbf{z}e^{j\phi}) - \frac{m}{k} 2\pi l(\mathbf{z}e^{j\phi}) - \frac{m}{k} \arg \mu_k(\mathbf{z}e^{j\phi}) \end{aligned} \quad (36)$$

$$\begin{aligned} &= \arg \mu_m(\mathbf{z}) + m\phi - \frac{m}{k} 2\pi l(\mathbf{z}e^{j\phi}) \\ &\quad - \frac{m}{k} \left(\arg \mu_k(\mathbf{z}) + k\phi + 2\pi \hat{l} \right) \end{aligned} \quad (37)$$

$$= \arg \mu_m(\mathbf{z}) - \frac{m}{k} \arg \mu_k(\mathbf{z}) - \frac{m}{k} 2\pi \left(l(\mathbf{z}e^{j\phi}) + \hat{l} \right) \quad (38)$$

$$= \nu(\mathbf{z}, l(\mathbf{z}e^{j\phi}) + \hat{l}), \quad (39)$$

where (36) uses (33), (37) uses (30), (38) are simple manipulations, and (39) uses (33) again. From (39) and our choice (34) applied to $\mathbf{z}e^{j\phi}$, we see that $\nu(\mathbf{z}e^{j\phi}, l(\mathbf{z}e^{j\phi})) = \nu(\mathbf{z}, l(\mathbf{z}e^{j\phi}) + \hat{l}) \in I$. Due to the fact that I is fixed, *i.e.*, that it does not depend on the shape \mathbf{z} , and due to the uniqueness (mod k) of the solution in (34) with respect to l , $l(\mathbf{z}e^{j\phi}) + \hat{l}$ must be equal to $l(\mathbf{z})$ (mod k). Replacing this concluding equality into (35) and using the definition of $\theta(\mathbf{z})$ in (29), property (21) is immediately obtained.

We call our method Principal Moment Analysis (PMA).

6.3 Dealing with Rotational Symmetry

We now deal with the case when there are no coprime k and m such that $\mu_k \neq 0$ and $\mu_m \neq 0$, in which case it is impossible to apply the method just described directly. Let γ be the greatest common divisor with respect to all the orders of non-zero PMs, *i.e.*, $\gamma = \gcd(K)$, $K = \{k \in \mathbb{Z}^+ : \mu_k \neq 0\}$. When there are no k and m coprime such that $\mu_k \neq 0$ and $\mu_m \neq 0$, we have $\gamma > 1$. This is equivalent to the existence of a γ -fold rotational symmetry, *i.e.*, that the shape is invariant to rotations of $2\pi/\gamma$. A simple way to derive this equivalence is to use the fact that, as shown in Section 3, the PMs are the coefficients of the Fourier series of the restriction $h(\mathbf{z}, \theta)$ in (9) of the ANSIG to the unit circle: since the rotation of a shape propagates into its ANSIG (Rodrigues et al, 2008a), γ -fold rotationally symmetric shapes lead to $h(\mathbf{z}, \theta)$ with period $2\pi/\gamma$, whose non-zero Fourier series coefficients will only occur at multiples of γ ; conversely, if the Fourier series coefficients only occur at multiples of γ , $h(\mathbf{z}, \theta)$ has period $2\pi/\gamma$ and, as such, using the same propagation property, the shape is γ -fold rotationally symmetric. In this case, all normalization angles of the form

$$\theta = \theta_0 + \hat{k} 2\pi/\gamma, \hat{k} \in \mathbb{Z} \quad (40)$$

lead to the same normalized shape. Hence, to compute a normalization angle, it suffices to compute the Fourier series coefficients of the function $h(\mathbf{z}, \theta/\gamma)$ instead of the ones of $h(\mathbf{z}, \theta)$ (in the variable θ), then to use these coefficients

(*i.e.*, the PMs) in the PMA as described in the previous subsection, obtaining an angle θ'_0 , and, finally, to invert the expanding effect of h through the contraction of θ'_0 , *i.e.*, assign $\theta_0 = \theta'_0/\gamma$. Any θ in (40) can then be used.

In terms of the PMs, it is easy to see through the properties of the Fourier series (Oppenheim et al, 1996) that this procedure is simply equivalent to using the PMs of orders γk and γm instead of the original ones of orders k and m , respectively, and then contracting the resulting angle. Finally, a last equivalent method is to compute the PMs (2) or (26) directly from the “powered” shape vector $[z_1^\gamma, z_2^\gamma, \dots, z_N^\gamma]^T$ (equal to the spectrum of $h(\mathbf{z}, \theta/\gamma)$ up to a real positive scaling factor), apply the PMA and contract the result.

6.4 Improving Robustness

Until now, we presented a theoretical proof for the correctness of PMA to unambiguously compute the orientation of arbitrary shapes using a pair of moments. Since in practice it is also important to obtain robustness to noise, we now describe how to improve the robustness of PMA by using a larger set of PMs. In fact, PMA can be used with any pair of coprime indices (k, m) , provided that $\mu_k \neq 0$ and $\mu_m \neq 0$. In order to improve robustness, we integrate the contributions of several pairs $(k_1, m_1), (k_2, m_2), \dots, (k_M, m_M)$, by computing pairwise estimates $\theta_i(\mathbf{z}), i = 1, 2, \dots, M$, and defining a robust

normalization angle $\theta(\mathbf{z})$ as the (angular) weighted average of them:

$$\theta(\mathbf{z}) = \arg \sum_{i=1}^M w_i e^{j\theta_i(\mathbf{z})}. \quad (41)$$

The reason for the angular average is its ability to deal with angles close to the region of circular discontinuity. For instance, we want the average of 1° and 359° to be $\arg(\exp(j 1^\circ) + \exp(j 359^\circ)) = 0^\circ$, not $(1^\circ + 359^\circ)/2 = 180^\circ$. The proof that property (21) holds for the robust normalization angle defined in (41) is straightforward:

$$\begin{aligned} \theta(\mathbf{z}e^{j\phi}) &= \arg \sum_{i=1}^M w_i e^{j\theta_i(\mathbf{z}e^{j\phi})} \\ &= \arg \sum_{i=1}^M w_i e^{j(\theta_i(\mathbf{z})+\phi)} \\ &= \arg \sum_{i=1}^M w_i e^{j\theta_i(\mathbf{z})} + \phi \\ &= \theta(\mathbf{z}) + \phi, \end{aligned} \quad (42)$$

$$(43)$$

where (42) uses the fact that the individual $\theta_i(\mathbf{z})$ satisfy (21) and (43) uses definition (41).

6.5 Implementation Details

In practice, PMA was implemented the following way: we start by computing the first 20 power sums in (26) and normalizing their magnitude by dividing

μ_k by its absolute value raised to k , *i.e.*, by using $\hat{\mu}_k = \mu_k/|\mu_k|^k$ instead of μ_k . This normalization aims to cancel the growth of the magnitude of the power sums with k . Note that alternatives to this normalization include the one discussed in Appendix A or the usage of the PMs as defined (2). Then, we detect non-zero PMs by thresholding their magnitude, *i.e.*, we only use $\hat{\mu}_k$ if $|\hat{\mu}_k|$ is above a threshold, say, 10^{-3} . In what respects to experimental tuning, our method only requires dealing with the two parameters just referred. Finally, we compute the greatest common divisor of all the indices corresponding to non-zero PMs and run the algorithm described in the previous section.

In the second step of the algorithm (the core of the method), for each non-zero moment $\hat{\mu}_k$, we find the smallest index m that is coprime with k (the first step described above guarantees that this index always exists). Then, we search for $l(\mathbf{z}) \in \{0, 1, 2, \dots, k-1\}$ until we find $\nu(\mathbf{z}, l(\mathbf{z})) \bmod 2\pi \in I$ (we choose $I = (-\pi/k, \pi/k]$). The value found for $l(\mathbf{z})$ is used to compute the pairwise estimates $\theta_i(\mathbf{z})$, which are then averaged using weights given by $w_i = |\hat{\mu}_{k_i} \hat{\mu}_{m_i}|$. The rationale for the choice of these weights is that PMs with larger magnitude have an argument less sensitive to noise. Note that the PMs used in the weights are the normalized ones, being this normalization thus also important to avoid over-weighting angles $\theta_i(\mathbf{z})$ computed from PMs of large order.

7 Extension to Grey-level Images

The algorithm presented in the previous section computes a unique orientation $\theta(\mathbf{z})$ for an arbitrary set of landmarks $\{z_1, z_2, \dots, z_N\}$, satisfying property (21). We now generalize the concept to compute a unique orientation of a continuous image $g(x, y)$. For that purpose, generalize the moments of (26) to the equivalent ones of Abu-Mostafa and Psaltis (1985), *i.e.*, make $\mu_k(g) = C_{k0}(g)$. From (4), the PMs of the grey-level image $g(x, y)$ are thus

$$\mu_k(g) = \int \int_{-\infty}^{+\infty} (x + jy)^k g(x, y) dx dy, \quad (44)$$

with $k \in \{0, 1, 2, \dots\}$. In what respects to representation, the generalization to continuous images loses completeness: in opposition to the case of a set of points, discussed in Sections 2 and 3, the PMs in (44) do not determine the image $g(x, y)$ univocally. An immediate way to conclude that is focusing on radial images. Start by rewriting (44) in polar coordinates:

$$\mu_k(g) = \int_{-\pi}^{\pi} \int_0^{\infty} r^k e^{jk\theta} g(r \cos \theta, r \sin \theta) r dr d\theta. \quad (45)$$

Now, for the radial image $g(r \cos \theta, r \sin \theta) = R(r)$, from (45), we easily get $\mu_k(g) = 0$, for $k \in \{1, 2, 3, \dots\}$, and $\mu_0(g) = 2\pi \int_0^{\infty} R(r) r dr$. As this integral does not define the function $R(r)$ univocally³, the PMs do not determine

³For example, the functions $R_1(r) = H(r) - H(r - \sqrt{2/3})$ and $R_2(r) = r(H(r) - H(r - 1))$, where $H(\cdot)$ denotes the Heaviside step function, lead to the same value for the moment $\mu_0 = 2\pi/3$.

$g(x, y)$.

Naturally, the lack of completeness just referred does not impede the extension of PMA, our rotational normalization algorithm presented in Section 6, for continuous images. A simple way to derive this extension is by using the derivations in Sections 4 and 6 with new definitions of the objects at hand, *i.e.*, with moments and rotations of images instead of shapes. The symbol \mathbf{z} is now interpreted as an image $g(x, y)$ and $\mathbf{z}e^{j\phi}$ is interpreted as the image that results from the (counterclockwise) rotation of \mathbf{z} by the angle ϕ . With these definitions, it is trivial to show that it still holds $[\mathbf{z}e^{j\phi}]e^{j\psi} = \mathbf{z}e^{j(\phi+\psi)} = [\mathbf{z}e^{j\psi}]e^{j\phi}$ (because image rotation is associative and commutative), thus the derivations in Section 4 remain valid for the interpretation in terms of continuous images.

As far as the derivation of the PMA algorithm in Section 6 is concerned, the reader should note that it is entirely based on the property $\mu_k(\mathbf{z}e^{j\phi}) = \mu_k(\mathbf{z})e^{jk\phi}$. We have to show that this property extends to the interpretation in terms of grey-level images, *i.e.*, that the k^{th} -order PM of an image rotated by ϕ equals to the product of the k^{th} -order PM of the original image by $e^{jk\phi}$ (notice that the multiplication of μ_k by $e^{jk\phi}$ is an ordinary one, not a rotation, since μ_k is a complex number, not an image). Using the definition

of PMs in (45), the desired property is immediate:

$$\begin{aligned}
\mu_k(\mathbf{z}e^{j\phi}) &= \int_{-\pi}^{\pi} \int_0^{\infty} r^k e^{jk\theta} g(r \cos(\theta-\phi), r \sin(\theta-\phi)) r dr d\theta \\
&= \int_{-\pi}^{\pi} \int_0^{\infty} r^k e^{jk(\theta+\phi)} g(r \cos \theta, r \sin \theta) r dr d\theta \\
&= e^{jk\phi} \int_{-\pi}^{\pi} \int_0^{\infty} r^k e^{jk\theta} g(r \cos \theta, r \sin \theta) r dr d\theta \\
&= e^{jk\phi} \mu_k(\mathbf{z}) .
\end{aligned}$$

Having shown how to extend PMA to grey-level images, we end this section by emphasizing that some care must be taken with the claims about the evidence of rotational symmetry and the universality of the algorithm. When dealing with sets of landmarks, as derived in Section 6, the moments μ_k are nonzero only for indices k multiples of $\gamma > 1$ if and only if the shape is γ -fold rotationally symmetric. However, this equivalence is not fulfilled for the case of continuous grey-level images. To get insight on what happens in this case, start by interpreting the moments in (45) in terms of the Fourier series of periodic signals obtained by circularly slicing the image. In particular, the k^{th} -order PM of the polar coordinate image $f(r, \theta) \stackrel{\text{def}}{=} g(r \cos \theta, r \sin \theta)$ can be written as

$$\mu_k = \int_0^{\infty} r^{k+1} F(r, k) dr , \quad (46)$$

where

$$F(r, k) = \int_{-\pi}^{\pi} f(r, \theta) e^{jk\theta} d\theta \quad (47)$$

are the coefficients of the Fourier series of the 2π -periodic (in θ) signal $f(r, \theta)$. Naturally, an image $f(r, \theta)$ is γ -fold rotationally symmetric if and only if it is $2\pi/\gamma$ -periodic in θ , for all $r > 0$. Hence, the coefficients $F(r, k)$ of its Fourier series are nonzero only for k multiple of γ if and only if there is γ -fold symmetry in $f(r, \theta)$ (see, *e.g.*, Oppenheim et al, 1996). In turn, the nonzero moments μ_k in (46) only occur for k multiple of γ in this case and PMA can be used as described in Subsection 6.3. However, this last statement is not an equivalence, since nonzero coefficients $F(r, k)$ may be destroyed by the integration in (46). In fact, as we detail in Appendix B, there exist very particular grey-level images that are not rotationally symmetric but have moments μ_k that are nonzero only for indices k multiples of $\gamma > 1$. There even exist images with a single non-zero moment (other than the radial ones referred in the first paragraph of this section; these last ones are not normalizable in what respects to orientation). Naturally, as any other method based on these moments, PMA fails to process these images.

8 Experiments

We now describe experiments. The following subsections focus on illustrating the compactness of the PM-based shape representation (Subsection 8.1), its usage in classification (Subsection 8.2), and the results of PMA for rotational normalization of shapes (Subsection 8.3) and grey-level images (Subsection 8.4).

8.1 PMs for Shape Representation

In this subsection, we focus on showing that the computational saving that arises from using our PMs does not degrade performance when compared with the discriminative ANSIG, the densely sampled signature introduced by Rodrigues et al (2008a). We illustrate this point with the shape shown in Fig. 8, which is described by 7 landmarks. The plot in Fig. 9 shows the magnitude of the PMs $\{M_k\}$ of this shape. Proceeding as described in Section 3, we obtain the required number of PMs for this shape, $k_B = 6$. As easily perceived from Fig. 9, the magnitude of the 6th PM is very small, indicating that the first 6, $\{M_k, 0 \leq k \leq 5\}$, containing the majority of the energy, adequately describe the shape. Since the shape was pre-processed as in (3), we obtain $M_1 = 0$ and $M_0 = 1$ (this last PM is not represented in the plot).

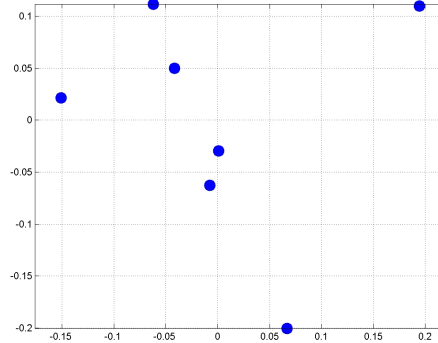


Figure 8: A 2D shape described by a small number of landmarks.

In Fig. 10, we represent, with solid lines, the magnitude and phase of

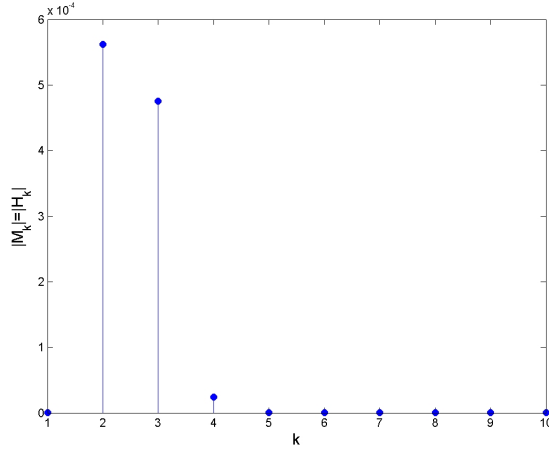


Figure 9: Magnitude of the PMs of the shape in Fig. 8 (or of the coefficients of the Fourier series of its analytic signature).

the densely sampled ANSIG of the shape in Fig. 8. As shown in Section 3, the coefficients $\{H_k\}$ of the Fourier series of this periodic complex signal are given by the PMs of the shape. To verify this in practice, we computed the *Fast Fourier Transform* (FFT) of the vector collecting the dense sampling of one period of the ANSIG, since it is straightforward to derive that this FFT is equal to $\{H_k\}$ multiplied by the number of samples (Oppenheim et al, 1996, 1999). As expected, we concluded that the Fourier series coefficients $\{H_k\}$ coincide with the PMs $\{M_k\}$, whose magnitude is represented in Fig. 9.

In the plots of Fig. 10, we also compare the densely sampled ANSIG with a signature obtained by interpolating our very compact representation. As derived in Section 3, the required number $k_B = 6$ of coefficients needed to represent the shape can be interpreted either as the minimum number of PMs or the minimum number of samples of the ANSIG. In Fig. 10, we

represent these $k_B = 6$ samples with stars, showing how much sparser this representation is when compared with the densely sampled ANSIG. Finally, we superimpose, represented by dashed lines, the reconstruction obtained by interpolating this sparse set as described in Section 3. We see that the lines of both plots are visually indistinguishable, showing that the PMs are adequate to represent the ANSIG and, consequently, the underlying shape. To illustrate what happens when using less samples than the minimum required by our study, we repeat the procedure by interpolating from 4 samples, obtaining Fig. 11. As easily seen, the reconstructed signature differs from the dense sampled one. Note, nevertheless, that there is no guarantee that our bound k_B is tight.

8.2 PMs for Shape Classification

In our experiments, the behavior illustrated in the previous subsection was observed in general, *i.e.*, a set of k_B PMs always suffices to accurately describe the densely sampled ANSIG. Although this is enough to guarantee that the same results are obtained when classifying shapes described by either their small sets of PMs or their dense ANSIGs, we also verified this directly. In particular, we generated noisy versions of prototype shapes and classified them by using 1-NN, *i.e.*, by just selecting the prototype that had most similar description. The number of PMs k_B ranged from 15 to 22, thus our descriptions are much shorter than the vectors of 512 ANSIG samples used in Rodrigues et al (2008a). We performed hundreds of tests for each

shape, obtaining the same performance (100% correct classifications, except for shapes that are visually indistinguishable), for both the densely sampled ANSIG and the PMs. Since the ANSIG was extensively demonstrated in shape-based classification of real images (Rodrigues et al, 2008a,b) and we have shown that the PMs have similar behavior, we do not report here other experiments on PM-based shape classification.

8.3 PMA for Shape Normalization

Although in the paper we have theoretically proven the correctness of PMA for shape normalization, *i.e.*, that it succeeds in unambiguously computing a unique orientation for any shape, the impact of PMA in shape-based recognition applications is also determined by the sensitivity to the noise, since observations of similar shapes must originate similar normalization angles. In this subsection, we illustrate that PMA is able to deal with these situations.

We start by illustrating that our method disambiguates the direction of the principal axis of generic shapes, *i.e.*, shapes without rotational symmetry. We used noisy versions of the shape in Fig. 1, see examples on the left column of Fig. 12, and computed the correspondent normalization angles using PMA. The right column of Fig. 12 shows the resultant shapes, *i.e.*, the rotationally normalized versions of the corresponding shapes on the left. We see that, regardless of the noise, all the normalized shapes exhibit similar orientation.

Although the examples in Fig. 12 illustrate the disambiguation of direction, the accuracy of the estimates of the normalization angle is better

evaluated by contrasting them with the ground truth. We thus performed experiments by rotating noisy shapes according to a known angle, ranging from $-\pi$ to π , and then estimating the orientation. The plot in Fig. 13 summarizes the results. We see that the estimates obtained by using PMA are very close to the correspondent true orientation, for all values of the rotation angle. For illustration purposes, the plot also shows the results obtained by using PCA, which naturally exhibit the directional ambiguity discussed in Section 1 and illustrated in Fig. 1.

In a similar way, we now illustrate that PMA also deals with noisy observations of rotationally symmetric shapes. We used noisy versions of a three-fold rotationally symmetric shape that would be impossible to orient using PCA, see the left column of Fig. 14, and normalized their orientations using PMA, obtaining the visually correct results on the right column of Fig. 14. By proceeding in a similar way as described above, we contrasted the estimates with their ground truth, obtaining the plot in Fig 15. Note that, in this case, there are three values for the true rotation angle, corresponding to three consistent orientations, due to the three-fold symmetry of the shape, see expression (40). Since the shape is rotationally symmetric, PCA is useless for the determination of an orientation, providing results only determined by the noise. In opposition, the plot in Fig 15 shows that the estimates obtained through PMA are robust to noise.

To illustrate the robustness of PMA to the shape sampling density, we used Japanese characters. We synthesized corrupted versions of those char-

acters by removing up to 95% of the shape points, obtaining shape vectors of the form of expression (1) but of very distinct cardinality. We then processed these vectors by using PMA. In Fig. 16, we single out five instances of a specific character to illustrate the consistent orientations obtained for all the corrupted characters.

8.4 PMA for Normalization of Grey-level Images

Finally, we illustrate the usage of PMA to rotationally normalize grey-level images. We used real images, consisting of photos of trademark logos, see examples in the left columns of Figs. 17-19. By proceeding as described in Section 7, we used PMA to normalize the orientation of the photographed logos. In spite of geometric distortions (*e.g.*, perspective, radial) and other intensity disturbances in the images, we got consistent results, as illustrated by the oriented versions of the logos in the right columns of Figs. 17-19. Note that these examples correspond to particularly challenging grey-level images, due to approximate rotational symmetry.

Since we detailed several situations where current methods fail, see Section 5, we do not report here experimental results obtained with those algorithms. In fact, it would be easy to produce examples where estimates obtained through PMA would be much more accurate than those obtained by using other methods (just imagine using shapes similar to the ones in Figs. 5, 6, or 7). However, we found it would be more informative to present the discussion in Section 5 regarding the core limitations of those methods,

i.e., to show how they attempt to use information that is not available in all shapes, than to blindly report sample experiments to support our approach.

9 Conclusion

We proposed to represent 2D shapes, *i.e.*, sets of unlabeled points or landmarks, via particular complex moments that we call *Principal Moments* (PMs). This representation is complete and we show it is compact, in the sense that the number of PMs needed to discriminate between shapes is small (and dependent on their complexity). We further presented a new method, *Principal Moments Analysis* (PMA) to unambiguously compute a unique orientation for arbitrary 2D shapes. This enables performing rotational normalization, thus obtaining maximally invariant (*i.e.*, complete) representations for 2D shapes. We finalized by extending PMA to the normalization of grey-level images. Besides theoretically sound, PMA results are robust to noise.

Acknowledgements

This work was partially supported by FET, within the EU-FP7, under SIMBAD project (contract 213250), and FCT, under ISR/IST plurianual funding (POSC program, FEDER) and grant MODI-PTDC/EEA-ACR /72201/2006.

A On the Normalization of Power Sums

In this appendix, we derive an expression for the expected growth of the magnitude of the power sum

$$\mu_k = \sum_{n=1}^N z_n^k, \quad k \in \{1, 2, 3, \dots\}, \quad (48)$$

under reasonable assumptions for the set of 2D points $\{z_n\}$.

Let $\{z_n, n = 1, 2, \dots, N\}$, be samples of a complex random variable (r.v.). Due to the common pre-processing of centering the shape (3), we assume this r.v. is zero mean, *i.e.*, $E\{z_n\} = 0, \forall_n$. We also assume the likelihood of each direction is the same, *i.e.*, the angle $\arg(z_n)$ is uniform in $[0, 2\pi)$ and independent of the absolute value $|z_n|$. Under these assumptions, we obtain

$$\begin{aligned} E\{z_n^k\} &= E\{|z|^k e^{jk \arg(z_n)}\} \\ &= E\{|z|^k\} E\{e^{jk \arg(z_n)}\} \end{aligned} \quad (49)$$

$$= 0, \quad (50)$$

where (49) is due to the independence between $|z_n|$ and $\arg(z_n)$ and (50) is due to the uniformity of $\arg(z_n)$. Using this result, we conclude that the mean value of the power sums is zero:

$$E\{\mu_k\} = E\left\{\sum_{n=1}^N z_n^k\right\} = \sum_{n=1}^N E\{z_n^k\} = 0. \quad (51)$$

The issue we address in the sequel is the expected grow of $|\mu_k|$.

We start by expressing $E\{|\mu_k|^2\}$ in terms of a moment of the real r.v. $|z|$, through the chain of equalities

$$\begin{aligned}
E\{|\mu_k|^2\} &= E\{\mu_k \mu_k^*\} \\
&= E\left\{\sum_{n=1}^N z_n^k \sum_{m=1}^N z_m^{k*}\right\} \\
&= \sum_{n,m=1}^N E\{(z_n z_m^*)^k\} \\
&= \sum_{n=1}^N E\{(z_n z_n^*)^k\} \tag{52}
\end{aligned}$$

$$\begin{aligned}
&= \sum_{n=1}^N E\{|z_n|^{2k}\} \\
&= NE\{|z_n|^{2k}\}, \tag{53}
\end{aligned}$$

where $*$ denotes the complex conjugate and (52) is due to the independence between z_n and z_m for $n \neq m$. Expression (53) states that $E\{|\mu_k|^2\}$ is proportional to the moment of order $2k$ of the real r.v. $|z|$ (this type of moments is often referred as *raw moments*, to emphasize that the corresponding r.v. is not zero mean, as it is obviously the case of $|z|$).

Depending on the probability density function (p.d.f.) of the r.v. z , we obtain different growing rates for $|\mu_k|$. For example, if the p.d.f. of z , besides being circularly symmetric on the complex plane, is Gaussian, *i.e.*, if z is a 2D Gaussian r.v. with co-variance proportional to the identity matrix, its absolute $|z|$ is a Rayleigh r.v. (see, *e.g.*, Papoulis, 1991). The p^{th} -order raw

moment of a Rayleigh r.v. is given by

$$\mathcal{M}_p = \text{E} \{ |z_n|^p \} = \sigma^p 2^{p/2} \Gamma(1 + p/2), \quad (54)$$

where

$$\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt$$

is the *Gamma function*, which, for an integer argument, is given by

$$\Gamma(q) = (q - 1)! \quad (55)$$

(see Papoulis, 1991). From (53), (54), and (55), we finally get

$$\begin{aligned} \text{E}\{|\mu_k|^2\} &= N \mathcal{M}_{2k} \\ &= N \sigma^{2k} 2^k \Gamma(1 + k) \\ &= N (2\sigma^2)^k k!. \end{aligned} \quad (56)$$

Thus, for shapes respecting our assumptions, the moment μ_k should be normalized according to the square root of (56), which, using Stirling's approximation (Paris and Kaminsky, 2001), can be shown to be $\sqrt{\text{E}\{|\mu_k|^2\}} = o(k!)$.

B Grey-level Images That Make PMA Fail

In this appendix we discuss the conditions under which grey-level images have only a few (or even a single) nonzero moments, preventing PMA to work as desired. As anticipated in Section 7, the limitations come from images that are not rotationally symmetric but “appear to be”, in the sense that their nonzero moments μ_k occur only for indices k multiples of a given $\gamma > 1$. These images $f(r, \theta)$ have nonzero Fourier series coefficients $F(r, k)$, given by (47), that are canceled out in μ_k by the integral in (46).

Consider, as an example, the image

$$f(r, \theta) = \frac{1}{\pi} R(r) (\cos \theta + \cos 2\theta) . \quad (57)$$

It is clear that the image $f(r, \theta)$ is not rotationally symmetric, since the fundamental period of $\cos \theta + \cos 2\theta$ is 2π . The Fourier series coefficients $F(r, k)$ in (47), for non-negative k , *i.e.*, the coefficients that determine the PMs in (46), are given by

$$F(r, k) = R(r) (\delta(k - 1) + \delta(k - 2)) , \quad (58)$$

where $\delta(\cdot)$ denotes the Dirac delta function. This expression is easily obtained from the Fourier series synthesis formula $f(r, \theta) = 1/2\pi \sum_{k=-\infty}^{+\infty} F(r, k) e^{-jk\theta}$ (see, *e.g.*, Oppenheim et al, 1996). From (46) and (58), we obtain the PMs

of the image $f(r, \theta)$:

$$\mu_k = \delta(k-1) \int_0^\infty r^2 R(r) dr + \delta(k-2) \int_0^\infty r^3 R(r) dr .$$

It is now clear that we can specify a function $R(r)$ such that only one PM is nonzero. For example, with

$$R(r) = H(r) - 2H(r-1) + H(r-\sqrt[3]{2}) , \quad (59)$$

where $H(\cdot)$ denotes the Heaviside step function, we obtain:

$$\begin{aligned} \mu_0 &= 0 \\ \mu_1 &= 0 \\ \mu_2 &= \frac{1 - \sqrt[3]{2}}{2} \\ \mu_3 &= 0 \\ \mu_4 &= 0 \\ \mu_5 &= 0 \\ \dots &= 0 . \end{aligned} \quad (60)$$

This shows that the PMA algorithm fails to process the image specified by (57) and (59), since, from (60), besides wrongly assuming the image is 2-fold rotationally symmetric, PMA would fruitlessly search for a pair of nonzero PMs. This grey-level image $f(r, \theta)$ is shown in Fig. 20.

We distinguish two cases where PMA applied to grey-level images fails, caused by the failure of two different parts of the algorithm. Namely, PMA fails when the image $f(r, \theta)$

- has a single nonzero PM, caused by
 - nonzero Fourier series coefficients $F(r, k)$ in (47) for more than one value of k but nonzero moments μ_k in (46) for a single value of k ;
 - nonzero $F(r, k)$ in (47) for a single $k \geq 1$ (the case $k = 0$ corresponds to a radial image, not normalizable in what respects to orientation);
- has zero PMs for all orders k not multiple of a given $\gamma > 1$ without being γ -fold rotationally symmetric, caused by
 - at least one Fourier series coefficient $F(r, k)$ in (47) is nonzero for k not multiple of γ but the nonzero moments μ_k in (46) occur only for k multiple of γ .

In the case of a single nonzero PM, the algorithm fails to find co-prime pairs, whereas in the case of nonzero PMs only for k multiple of a given $\gamma > 1$, the failure lies on the incorrect detection of a γ -fold symmetry. Grey-level images of these classes appear to be somewhat particular, as the example in Fig. 20 illustrates, thus we did not face any failure when processing real images of trademark logos, see examples in Figs. 17-19.

References

- Abu-Mostafa Y, Psaltis D (1985) Image Normalization by Complex Moments. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 7(1)
- Ahlfors L (1978) *Complex Analysis*. McGraw-Hill, New York, NY, USA
- Atallah M (1985) On Symmetry Detection. *IEEE Transactions on Computers* 34(7)
- Bartolini I, Ciaccia P, Patella M (2005) Warp: Accurate Retrieval of Shapes using Phase of Fourier Descriptors and Time Warping Distance. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 27(1):142–147
- Besl P, McKay N (1992) A Method for Registration of 3-D Shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 14(2)
- Chauang G, Kuo C (1996) Wavelet Descriptor of Planar Curves: Theory and Applications. *IEEE Transactions on Image Processing* 5:56–70
- Derrode S, Ghorbel F (2004) Shape Analysis and Symmetry Detection in Gray-Level Objects Using the Analytical Fourier-Mellin Representation. *Signal Processing* 84(1)
- Ha V, Moura J (2003) Efficient 2D Shape Orientation. In: *Proc. of IEEE Int. Conf. on Image Processing, Barcelona, Spain*

- Ha V, Moura J (2005) Robust Reorientation of 2D Shapes Using the Orientation Index. In: Proc. of IEEE Int. Conf. on Acoustics, Speech, and Signal Processing, Philadelphia, PA, USA
- Hu M (1962) Visual Pattern Recognition by Moment Invariants. IRE Transactions on Information Theory 8:179–187
- Jebara T (2003) Images as Bags of Pixels. In: Proc. of IEEE Int. Conf. on Computer Vision, Nice, France
- Kanatani K (1990) Group Theoretical Methods in Image Understanding. Springer Verlag, Berlin, Germany
- Kendall D, Barden D, Carne T, Le H (1999) Shape and Shape Theory. John Wiley and Sons
- Khotanzad A, Hong Y (1990) Invariant Image Recognition by Zernike Moments. IEEE Transactions on Pattern Analysis and Machine Intelligence 12:489–497
- Kondor R (2008) Group theoretical methods in machine learning. PhD thesis, Columbia University, New York, NY, USA
- Lin JC (1993) Universal Principal Axes: an Easy-to-Construct Tool Useful in Definig Shape Orientations for Almost Every Kind of Shape. Pattern Recognition 26(4)

- Lin JC, Tsai WH, Chen JA (1994) Detecting Number of Folds by a Simple Mathematical Property. *Pattern Recognition Letters* 15(10)
- Marola G (1989) On the Detection of Axes of Symmetry of Symmetric and Almost Symmetric Planar Images. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 11(6)
- McNeill G, Vijayakumar S (2006) Hierarchical Procrustes Matching for Shape Retrieval. In: *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition*, New York, NY, USA
- Mukundan R, Ong S, Lee P (2001) Image Analysis by Tchebichef Moments. *IEEE Transactions on Image Processing* 10:1357–1364
- Oppenheim A, Willsky A, Hamid S (1996) *Signals and Systems*, 2nd edn. Prentice Hall, Upper Saddle River, NJ, USA
- Oppenheim A, Schafer R, Buck J (1999) *Discrete-Time Signal Processing*. Prentice Hall, Upper Saddle River, NJ, USA
- Papoulis A (1991) *Probability, Random Variables, and Stochastic Processes*, 3rd edn. McGraw Hill, New York, NY, USA
- Paris R, Kaminsky D (2001) *Asymptotics and the Mellin-Barnes Integrals*. Cambridge University Press, Cambridge, UK
- Prasad V, Yegnanarayana B (2004) Finding Axes of Symmetry from Potential Fields. *IEEE Transactions on Image Processing* 13(12)

- Rodrigues J, Aguiar P, Xavier J (2008a) ANSIG – An Analytic Signature for Permutation-Invariant Two-Dimensional Shape Representation. In: Proc. of IEEE Conf. on Computer Vision and Pattern Recognition, Anchorage, AK, USA
- Rodrigues J, Xavier J, Aguiar P (2008b) Classification of unlabeled point sets using ANSIG. In: Proc. of IEEE Int. Conf. on Image Processing, San Diego, CA, USA
- Shen D, Ip H (1997) Generalized Affine Invariant Image Normalization. IEEE Transactions on Pattern Analysis and Machine Intelligence 19(5)
- Shen D, Ip H (1999) Symmetry Detection by Generalized Complex (GC) Moments: a Closed-Form Solution. IEEE Transactions on Pattern Analysis and Machine Intelligence 21(5)
- Teh CH, Chin R (1988) On Image Analysis by the Method of Moments. IEEE Transactions on Pattern Analysis and Machine Intelligence 10(4)
- Tsai W, Chou S (1991) Detection of Generalized Principal Axes in Rotationally Symmetric Shapes. Pattern Recognition 24(2)
- Zunic J, Kopanja L, Fieldsend J (2006) Notes on Shape Orientation Where the Standard Method Does Not Work. Pattern Recognition 39(5)

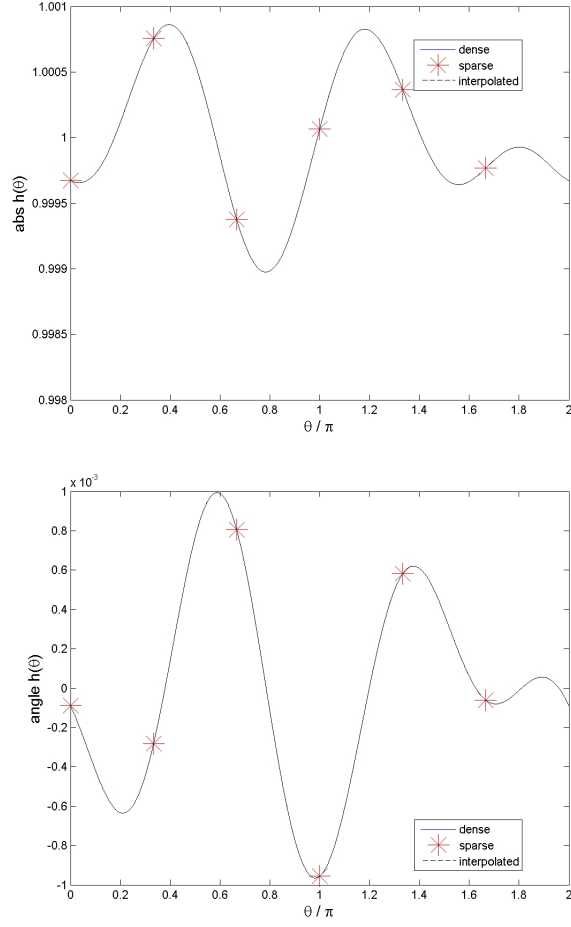


Figure 10: Magnitude and phase of the analytic signature of the shape in Fig. 8. The very sparse representation we propose leads to plots that result visually indistinguishable from those obtained by dense sampling (after reconstruction).

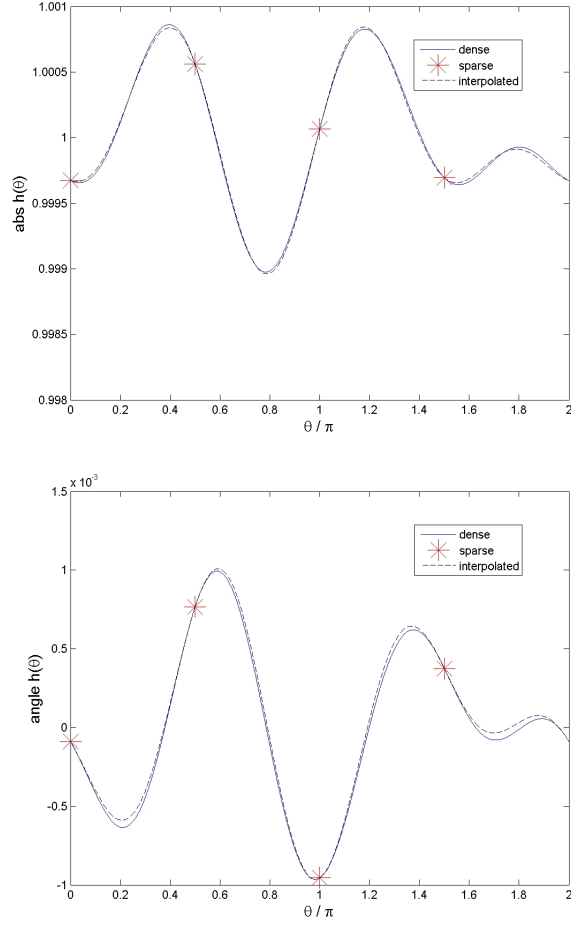


Figure 11: Same as in Fig. 10, now with a number of samples smaller than the one required by our study. Note how the interpolated signature now differs from the dense sampled one.

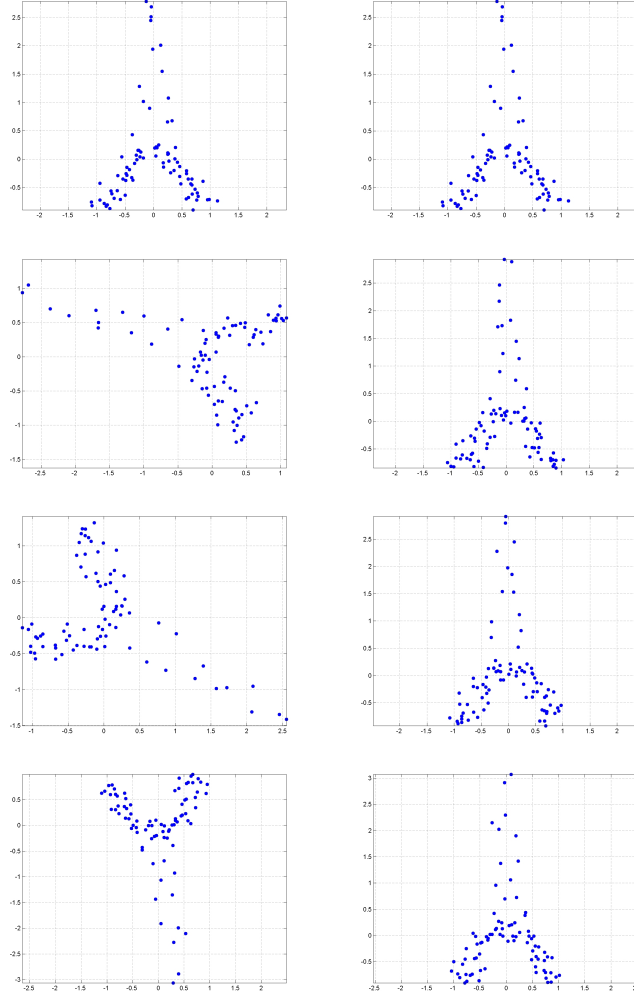


Figure 12: Examples of using PMA for computing the orientation of general shapes. Left: original shapes; right: PMA oriented shapes. Relative to the ambiguity illustrated in Fig. 1, note that PMA disambiguates the direction of the principal axis.

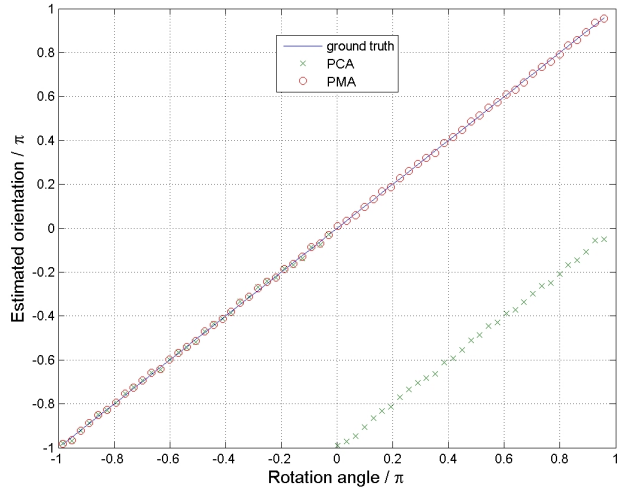


Figure 13: Comparison of PMA and PCA for general shapes.

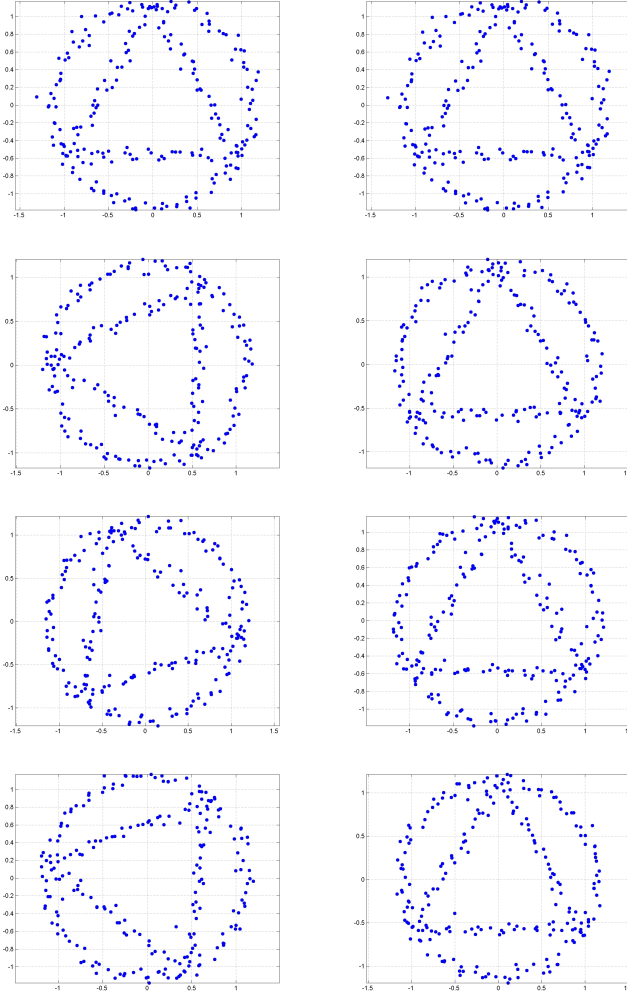


Figure 14: Examples of using PMA for computing the orientation of rotationally symmetric shapes. Left: original shapes; right: PMA oriented shapes. In spite of the absence of a principal axis, and the high level of noise, PMA provides consistent orientations.

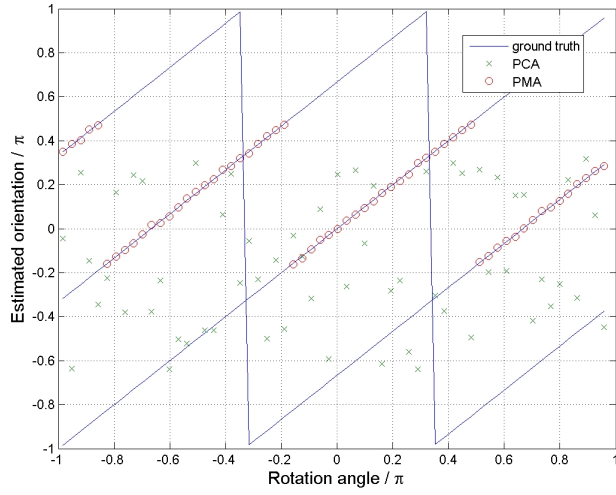


Figure 15: Comparison of PMA and PCA for rotationally symmetric shapes. In this case, with a three-fold symmetric shape, the ground true is represented by three lines corresponding to the three consistent orientations, *i.e.*, separated by $2\pi/3$.

猿 猿
猿 猿
猿 猿
猿 猿
猿 猿
猿 猿

Figure 16: Using PMA to normalize corrupted Japanese characters.

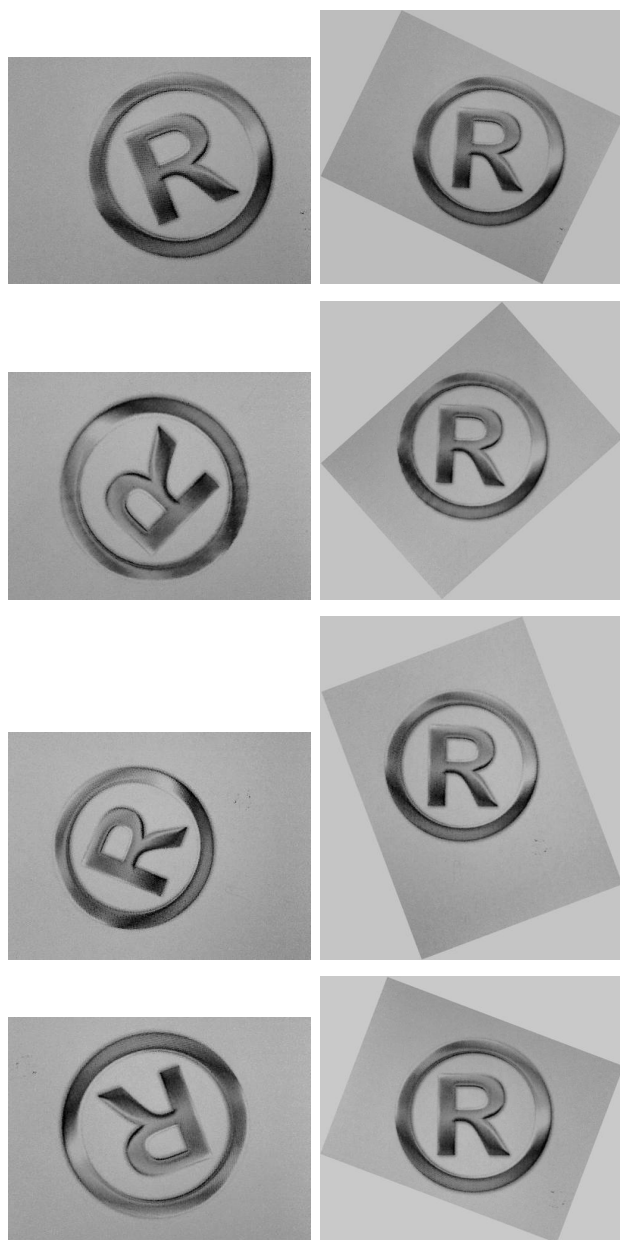


Figure 17: Using PMA with gray-scale photos of trademark logos.

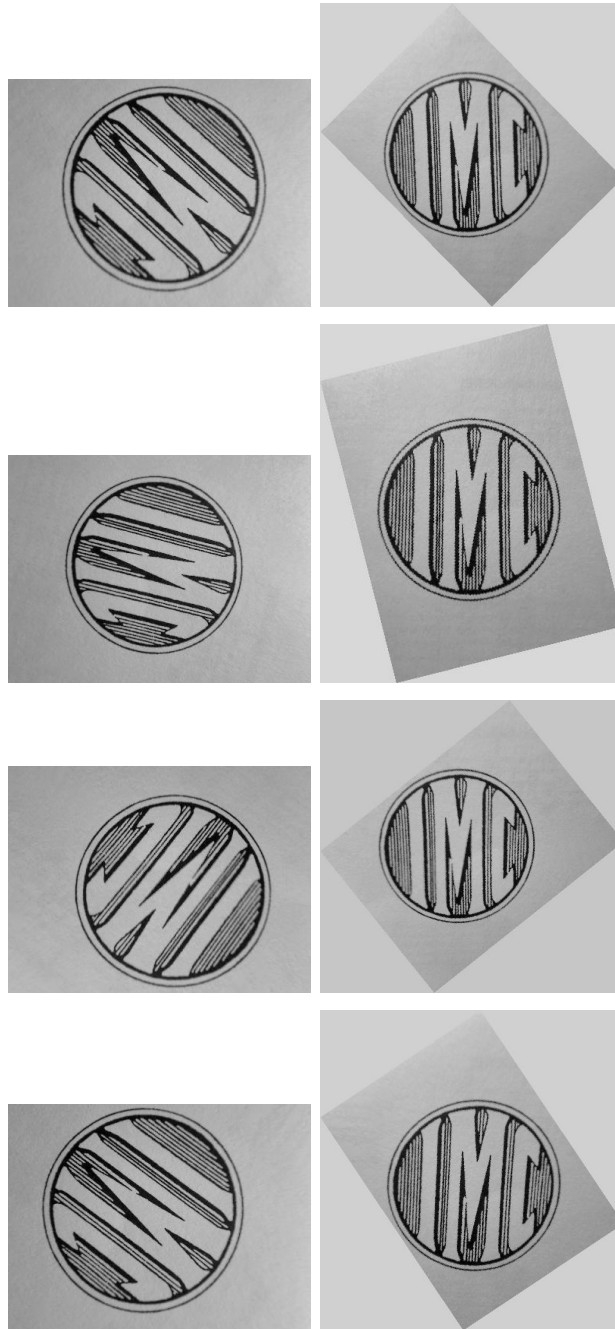


Figure 18: Using PMA with gray-scale photos of trademark logos.

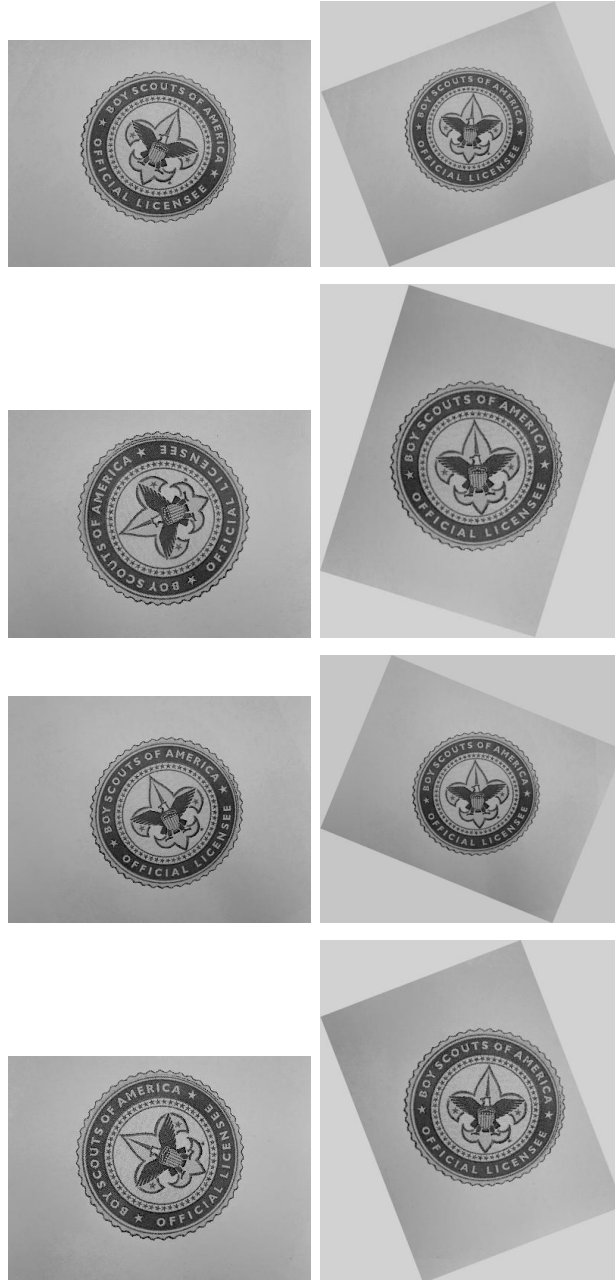


Figure 19: Using PMA with gray-scale photos of trademark logos.

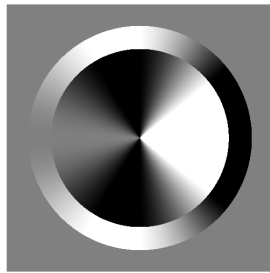


Figure 20: Example of a grey-level image for which PMA fails.